Modeling System Audit as a Sequential test with Discovery as a Failure Time Endpoint

Craig S Wright & Tanveer A Zia

[1] School of Computing and Mathematics
Charles Sturt University
Wagga Wagga, NSW, Australia
crwright@csu.edu.au, tzia@csu.edu.au

Abstract. Combining hazard models with SIR (Susceptible-Infected-Removed) epidemic modeling provides a means of calculating the optimal information systems audit strategy. Treating audit as a sequential test allows for the introduction of censoring techniques that enable the estimation of benefits from divergent audit strategies. This process can be used to gauge the economic benefits of these strategies in the selection of an optimal audit process designed to maximize the detection of compromised or malware infected hosts.

Keywords: Modeling, Hazard, non-homogeneous Poisson process (NHPP), failure intensity, SIR (Susceptible-Infected-Removed), epidemic modeling.

1 Introduction

Computer systems are modeled through periodic audit and monitoring activities. This complicates the standard failure and hazard models that are commonly deployed [11]. A system that is found to have been compromised by an attacker, infected by malware or simply suffering a critical but unexploited vulnerability generally leads to early intervention. This intervention ranges from system patching or reconfiguration to complete rebuilds and decommissioning.

Audits and reviews of computer systems usually follow a prescribed schedule in chronological time. This may be quarterly, annually or to any other set timeframe. Further, periodic reviews and analysis of systems in the form of operational maintenance activities also provide for a potential intervention and discovery of a potential system failure or existing compromise.

Using a combination of industry and organizational recurrence rates that are stipulated from a preceding failure and covariate history as derived from the individual organization introduces a rational foundation in modeling current event data. By denoting the number of incidents[1] within the organization as $\tilde{N}(t)$ by

[1] An incident as defined for the purposes of this paper is an event leading to the failure of the system. This can include a system compromise from an attacker or an infection process of malware (such as a scanning worm).

follow-up time t and $N(t)$ as the corresponding observed incidents in $(0,t]$ with regards to absolute continuous event times, the hazard or intensity process $\lambda(t)$ for the intervention time t using the covariate data $X(t)$ can be expressed as:

$$\lambda(t) = P\left[d\tilde{N}(t)\right] = 1\left[\tilde{N}(u), 0 \le u < t, X(t)\right] \quad (1)$$

Taking the assumption that the administrative and audit staff are not the direct cause of an incident, a point process $(T_1, T_2, T_3, ...)$ will usually be observed for the system[2] being examined. Due to censoring through the audit process, $N(t)$ can be greater then $\tilde{N}(t)$. Equation (9.1) has an assumption that only a single incident has occurred, that is, $\tilde{N}$ increments by units. Live systems can and do experience multiple incidents and compromises between detection events. Hence it is also necessary to model the mean increments in $\tilde{N}$ over time

$$d\Lambda(t) = E\left[\tilde{N}(t) \mid \tilde{N}(u), 0 \le u < t, X(t)\right] \quad (2)$$

with the cumulative intensity process Λ.

In the case of a continuous-time process with unit jumps, expressions (1) and (2) can be expressed as

$$\Lambda(t) = \int_0^t \lambda(u)\,du\,. \quad (3)$$

Independent censorship requires that $C \ge t$ [15]. This assumption of independent censorship allows the preceding covariate histories to be incorporated into the model. If we define $Y(t) = 1(0 < t \le C)$, it is now necessary that

$$E\left[dN(t) \mid N(u), Y(u); 0 \le u < t, X(t)\right] = Y(t)\Lambda(t) \quad (4)$$

for all times ($t \le C$) prior to the audit or review.

2 NHPP, Non-homogeneous Poisson Process

Poison processes have been used to model software [11] and systems failures [16], but these models are too simplistic and it is necessary to vary the intensity (rate) based on historical and other data in order to create accurate risk models for computer systems. The non-homogeneous Poisson process (NHPP) can be used to model a Poisson process with a variable intensity. In the special case when $\lambda(t)$ takes a

[2] A system is defined by an isolated and interactive grouping of computers and processes. This could be a collection of client and server hosts located at a specific location isolated by a common firewall.

constant value λ, the NHPP is reduced to a homogeneous Poisson process with intensity $\lambda(t)=\lambda$.

In the heterogeneous case, an NHPP with intensity $\lambda(t)$, the increment, $N_t - N_u, 0 \le u < t$ has a Poisson distribution with an intensity of $\lambda(t)=\int_u^t \lambda(x)\,dx$. Hence the distribution function of the incident discovery can be expressed as:

$$\begin{aligned} \Lambda_u &= 1 - P\left(N_{u+t} - N_t = 0\right) \\ &= 1 - \exp\left(-\int_u^{u+t} \lambda(x)\,dx\right) \\ &= 1 - \exp\left(-\int_0^t \lambda(u+v)\,dv\right) \end{aligned} \tag{5}$$

The NHPP format is better suited to information systems risk modeling then is the homogeneous Poisson Process as it can incorporate changes that occur over time across the industry.

This can also be modeled as the Poisson process with parameter λ, $\left(N_t^{(\lambda)}, t \ge 0\right)$, is the unique (in law) increasing right continuous process with independent time homogeneous increments. Each $t>0, N_t^{(\lambda)}$ has a Poisson distribution with rate λt. The process $\left(X_t^{(\lambda)}, t \ge 0\right)$ is also stationary with independent time increments.

With $t_0 = 0$ and $(t_1,...,t_n) \in \mathbb{R}_+^n$, $t_1 < t_2 < ... < t_n$ the r.v.'s $\left[\left(X_{t_1}^{(\lambda)}\right), \left(X_{t_2}^{(\lambda)} - X_{t_1}^{(\lambda)}\right), ..., \left(X_{t_n}^{(\lambda)} - X_{t_{n-1}}^{(\lambda)}\right)\right]$ are independent and for each $k = 1,...,n, \left(X_{t_n}^{(\lambda)} - X_{t_{n-1}}^{(\lambda)}\right)$ has the same distribution as $\left(X_{t_k - t_{k-1}}^{(\lambda)}\right)$.

The characteristic function of $\frac{1}{\sqrt{\lambda}}\left(X_{t_k - t_{k-1}}^{(\lambda)}\right)$ can be computed for any $\lambda_m \in \mathbb{R}$ as:

$$\begin{aligned} E\left[e^{\frac{i\lambda_m}{\sqrt{\lambda}}\left(X_{t_k - t_{k-1}}^{(\lambda)}\right)}\right] &= e^{\left(-i\lambda^{1/2}\lambda_m(t_k - t_{k-1}) - \lambda(t_k - t_{k-1})\right)} \sum_{n=0}^{\infty}\left(\frac{\lambda^n (t_k - t_{k-1})^n}{n!} e^{\left(\frac{i\lambda_m n}{\sqrt{\lambda}}\right)}\right) \\ &= e^{\left(-\lambda(t_k - t_{k-1})\left(1 - e^{\left(\frac{i\lambda_m}{\sqrt{\lambda}}\right)}\right) - i\lambda_m(t_k - t_{k-1})\sqrt{\lambda}\right)} \end{aligned} \tag{6}$$

as $\lambda \to \infty$ the expression converges to $Exp\left(\frac{-\lambda_m^2}{2}(t_k - t_{k-1})\right)$, which is the characteristic function of a Gaussian variable with variance $\sigma^2 = (t_k - t_{k-1})$.

3 Recurrent Events

In many cases, audit and review processes are limited in scope and may not form a complete report of the historical processes that have occurred on a system. The audit samples selected systems and does not check neighboring systems unless a failure is discovered early in the testing. In these instances, the primary interest resides in selected marginalized intensities that condition only on selected parts of the preceding histories. Some marginal intensity rates drop the preceding incident history all together

$$d\Lambda_m(t) = E\left[d\tilde{N}(t) \mid X(t)\right] \tag{7}$$

A common condition for the identification of Λ_m is that

$$E\left\{dN(t) \mid \left[Y(u); 0 \le u < t\right], X(t)\right\} = T(t)\, d\Lambda_m(t). \tag{8}$$

For (6) to be valid, censoring intensity cannot depend on the preceding incident history for the system $\left[N(u); 0 \le u < t\right]$. The process of randomly selecting systems to audit makes it unlikely that particularly problematic systems will be re-audited on all occasions. This would include the exclusion of targeting client systems that have been compromised several times in the past or which have suffered more than one incident in recent history. The result is that covariates that are functions of $\left[\tilde{N}(u); 0 \le u < t\right]$ will also have to be excluded from the conditioning event. Here

$$E\left[d\tilde{N}(u) \mid X(u)\right] = E\left[d\tilde{N}(u) \mid X(t)\right], \qquad \forall \qquad t \ge u. \tag{7}$$

When this occurs

$$\begin{aligned} E\left[d\tilde{N}(u) \mid X(u)\right] &= \int_0^t E\left[d\tilde{N}(u) \mid X(t)\right] \\ &= \int_0^t E\left[d\tilde{N}(u) \mid X(u)\right] \\ &= \Lambda_m(t) \end{aligned} \tag{9}$$

$\Lambda_m(t)$ models the expected number of incidents that have occurred in the system over $(0, t]$ as a function of $X(t)$.

4 Cox Intensity Models

Using a Cox-type model

$$d\Lambda(t) = d\Lambda_0(t) e^{\left[Z(t)'\beta\right]}, \tag{10}$$

with $Z(t)' = \left[Z_1(t), ..., Z_p(t)\right]$ having been created using functions of $X(t)$ and $\left[N(u); 0 \le u < t\right]$, inference differs little to univariate failure time data. The log-partial likelihood function, score statistic and the integral notation for the information matrix may be written respectively as

$$\ell(\beta) = \log L(\beta) = \sum_{i=1}^{n}\left[\int_0^\infty \left\{Z_i(t)'\beta - \log\left[S^{(0)}(\beta,t)\right]\right\} dN_i(t)\right] \tag{9}$$

$$U(\beta) = \frac{\partial \ell(\beta)}{\partial \beta} = \sum_{i=1}^{n}\left[\int_0^\infty \left\{Z_i(t) - \xi(\beta,t)\right\} dN_i(t)\right] \tag{10}$$

and

$$I(\beta) = \frac{-\partial^2 \ell(\beta)}{\partial\beta\partial\beta'} = \sum_{i=1}^{n}\left[\int_0^\infty \left\{V(\beta,t)\right\} dN_i(t)\right] \tag{11}$$

where,

$$S^{(j)}(\beta,t) = \sum_{i=1}^{n} Y_i(t) Z_i(t)^{\otimes j} e^{Z_i(t)'\beta}$$

$$\xi(\beta,t) = \frac{S^{(1)}(\beta,t)}{S^{(0)}(\beta,t)}$$

and

$$V(\beta,t) = \frac{S^2(\beta,t)}{S^0(\beta,t)} - \xi(\beta,t)^{\otimes 2}. \tag{11}$$

By defining $Z(t)$ in terms of fixed or external time varying covariates, (10) can be further defined by adding additional elements $1\left[N(t^-)=1\right]\gamma_1 + 1\left[N(t^-)=2\right]\gamma_2 + ...$ to $Z(t)'\beta$. This would allow the intensity to be altered by a multiplicative factor e^{γ_j} following the jth incident on an individual system when compared against another system without any incidents at the same point in time.

5 SIR (Susceptible-Infected-Removed) epidemic modeling of incidents during Audit

Allowing that a compromised or infected system remains infected for a random amount of time τ , the discovery of an incident by an auditor will be dependent on a combination of the extent of the sample tested during the audit[3] and the rate at which the incident impacts individual hosts. When a host in a system is infected, any neighboring hosts are attacked and infected at a rate r. The sample size selected in the audit is set as κ and the total number of hosts in the system being audited is defined by K were $\kappa \leq \mathrm{K}$. The time between audits (the censor time) is defined by C.

If $C \leq \tau$, an infected or compromised system will be undiscovered and attacking other hosts within the system when the audit occurs. At the end of the time τ , the system is removed as it is either 'dead' - that is decommissioned and reinstalled or it has been patched against the security vulnerability.

A NSW [17] random graph is obtained by investigating the neighboring systems in the SIR model. From this, the thresholds can be computed.

5.1 Calculations with a constant τ .

First, we shall consider the case where τ is a constant value, and without loss of generality scale time to make a constant. Start with letting p_k be the degree of distribution.

Starting with a single infected host in a system, we can compute the probability that j of k neighboring hosts will be infected is given by:

$$\hat{p}_j = \sum_{k=j}^{\infty} p_k \binom{k}{j} \left(1-e^{-r}\right)^j e^{-(k-j)r} \tag{12}$$

Setting μ is the mean of p then the mean of $\hat{p}$ is $\hat{\mu} = \mu\left(1-e^{-r}\right)$

With the network constructed as an NSW random graph, systems that are compromised in the first and subsequent iterations will each have k neighbors. The value k includes subsequently compromised machines and the host that compromised the existing system. The probability associated with these neighboring hosts is give by:

$$q_k = \frac{(k+1)\, p_{k+1}}{\mu} \qquad \forall \quad k \geq 0 \tag{13}$$

This allows us to calculate the probability that j neighboring hosts also become infected:

[3] We shall assume that the audit is effective and will uncover an incident if an infected host is reviewed.

$$\hat{q}_j = \sum_{k=j}^{\infty} q_k \binom{k}{j} \left(1-e^{-r}\right)^j e^{-(k-j)r} \quad (14)$$

Setting ν to represent the mean of q, we get the mean of $\hat{q}$,

$$\hat{\nu} = \nu\left(1-e^{-r}\right).$$

From this we see that for the attack or malware to propagate and infect other systems, we have to have;

$$\nu\left(1-e^{-r}\right). \quad (15)$$

Using this, we can calculate the probability that a particular attack or type of malware resulting in an outbreak (such as[4] []). Setting $T = 1-e^{-r}$ []^5 we get:

$$\begin{aligned} \hat{G}_0(z) = & \sum_{k=0}^{\infty}\sum_{j=0}^{\infty} p_k \binom{k}{j} T^j (1-T)^{k-j} z^j \\ = & \sum_{k=0}^{\infty} p_k \sum_{j=0}^{\infty} \binom{k}{j} (Tz)^j (1-T)^{k-j} \\ = & G_0\left(Tz+(1-T)\right). \end{aligned} \quad (16)$$

Similarly, we can prove that $\hat{G}_1(z) = G_1\left(Tz+(1-T)\right)$.

As such, we can state that the probability that an incident behaves as an epidemic is $1-\hat{G}_0(\xi)$ where $\hat{G}_1(\xi) = \xi$ is the smallest fixed point in $[0,1]$.

5.2 Calculations with a variable τ.

Next we consider the effects of a variable or random value of τ. This is the probability that a compromised system causes a compromise in its neighboring system is:

$$T = 1-\int_0^{\infty} dtP(\tau = t)e^{-rt}. \quad (17)$$

Again, T is the transmissibility factor.

Newman [17] asserted that the infection of neighbors was independent. This does not hold as valid for malware, but it gives a good approximation. Interactions in systems and the ability of software to rescan the same systems carry a degree of dependence. Here the time to compromise may be modeled exponentially with mean λ, such that $P(\tau = t) = e^{-\lambda t}$.

[4] http://www.securecomputing.net.au/News/81027,google-warns-of-web-malware-epidemic.aspx

[5] Newman NSW

From this we can see that the probability of a host not being compromised is,

$$1-T=\int_0^\infty e^{-\lambda t}e^{-rt}dt=\int_0^\infty e^{-(\lambda+r)t}dt$$

$$=\frac{1}{(\lambda+r)} \tag{18}$$

Likewise, the probability that n hosts in a system are not compromised is,

$$1-T=\int_0^\infty e^{-\lambda t}e^{-nrt}dt=\int_0^\infty e^{-(\lambda+nr)t}dt$$

$$=\frac{1}{(\lambda+nr)} \tag{19}$$

For cases where τ is not constant, Jensen's inequality [6] implies that for neighboring hosts, the probability of escaping compromise is positively correlated as

$$E\left(e^{-nrt}\right)>\left(Ee^{-rt}\right)^n \tag{20}$$

We can compute the expected number of systems that will be compromised by substituting r_k for p_k and q_k which gives us,

$$\hat{G}(z)=\int_0^\infty P(\tau=1)dt\sum_{j=0}^{\infty}z^j\sum_{j=0}^{\infty}r_k\binom{k}{j}\left(1-e^{-rt}\right)^j e^{-r(k-j)t} \tag{21}$$

if $r_0+r_1>1$, G is strictly convex as $\hat{\nu}=\nu T$, $\tilde{G}_i=G_i\left(1+(z-1)T\right)$.

6 Applications to audit and review

The first case where $C>\tau$ has a host in the system being discovered as having been infected or compromised before the audit. Here, the rate of infection r determines the chance of other systems being uncovered during an audit. If the time between audits exceeds the time to compromise the first host is insufficient for the incident to spread (i.e. $\tau\le C, C<r$), then only the initial host will have been compromised and this will be known prior to the audit.

If $C\le\tau$ a compromised host in the system will be undiscovered and attacking other hosts within the system when the audit occurs. At the end of the time τ, the system is found. As such, if $(C+c)\le\tau$ (where c is the average time taken to conduct an audit), a compromised host is discovered during the audit through a process independent of the audit.

The alternative scenario and that which is of most interest is where $(C+c)<\tau$. In this case, the incident will not be discovered independent of the audit. In this instance, we can calculate the probability that an auditor will discover a compromised system during the audit process. In this instance we have a time limited network

function which is coupled with a discovery process that is formulated using the Bayesian prior. That is we can calculate the probability that all systems are not compromised given a selected audit strategy that finds that none of the audited hosts have been compromised.

6.1 False Negatives in an Audit

False negatives result in an audit where an incorrectly reported result is supplied noting the organization as safe when it is not (i.e. no compromise was detected where hosts have been compromised). If we let A represent the condition where the organization has been compromised and let B represent the positive evidence of a compromise being reported

$$P\left(A \mid \bar{B}\right) = \frac{P\left(\bar{B} \mid A\right)P\left(A\right)}{P\left(\bar{B} \mid A\right)P\left(A\right) + P\left(\bar{B} \mid \bar{A}\right)P\left(\bar{A}\right)} \tag{22}$$

Here we can model the actual rate of compromise in the system, $P\left(A\right)$. Given a network compromise model (21) we can substitute the censored time

$$\hat{G}\left(z\right) = \int_0^C P\left(\tau \le C\right)dt \sum_{j=0}^{\infty} z^j \sum_{j=0}^{\infty} r_k \binom{k}{j} \left(1 - e^{-rt}\right)^j e^{-r(k-j)t} \tag{23}$$

From this we can see that the probability of a host not being compromised in the censor time is,

$$\begin{aligned} P\left(\bar{A}\right) = 1 - T &= \int_0^C e^{-\lambda t} e^{-rt} dt = \int_0^C e^{-(\lambda + r)t} dt \\ &= \frac{e^{-(\lambda + r)t}}{-\left(\lambda + r\right)}\Bigg|_0^C = \frac{e^{-(\lambda + r)C}}{-\left(\lambda + r\right)} + \frac{1}{\left(\lambda + r\right)} \\ &= \frac{1 - e^{-(\lambda + r)C}}{\left(\lambda + r\right)} \end{aligned} \tag{24}$$

Equation (24) also derives the probability of any single host being compromised between the audits

$$P\left(A\right) = 1 - \frac{1 - e^{-(\lambda + r)C}}{\left(\lambda + r\right)}. \tag{25}$$

Depending on whether τ is constant or varies; we can calculate the expected number of hosts that will be compromised in the period between audits as a fixed or variable function. In either event, each calculation is an exercise in Bayesian estimation of the type where a random sample is selected and the defects or failures are analyzed

$$f\left(x\right) = \binom{n}{x} p^x q^{n-x}. \tag{26}$$

This binomial distribution is simplified in the case of a false negative (no failures or x=0 from a sample of n hosts)

$$f(x) = \binom{n}{0} p^x q^{n-x} = p^x q^{n-x}. \quad (27)$$

Based on the types of systems, the audit periods can be selected to create an economically optimal choice. In this, using an equation that calculates the expected number of compromised or infected hosts within a censor time we can select either a fixed audit schedule, C = Constant, or vary C over the course of the system life in order to maximize the detection, C=C(t).

In conducting this exercise, the cost of the audit, and differences that occur would also need to be modeled. The required effort for an audit of 10 hosts in a 1 month period is not necessarily linearly related to the audit of 60 hosts in a 6 month period. In each case, the individual constraints faced by the selected organization also need to be incorporated.

Conclusion

These equations allow organizations to compare the deployed audit strategies against both their own historical data and that of third parties. In this manner, strategy can be formulated in order to optimize audits and system reviews in a manner that detects an incident in the most economical manner.

Modeling the failure rate of systems and the propagation rate of an attack, allows us to calculate an expected number of hosts that are anticipated to have been compromised in the time between an audit given a specified survival function or threat. Past data and comparisons from similar systems (such as survival data from DShield[6]) allow for the modeling of alternative systems where a reported number of events have been reported against those deployed.

Dependence, variation, randomness, and frailty add to the risk toolset of multivariate failure event analysis. Using frailty theory to model information system risk allows us to better predict risk and to more effectively allocate scarce resources through selecting the most economically viable targets to defend as well as choosing the optimal detection strategies. The properties of censoring-handling and frailty modeling have turned multivariate survival analysis into an exceptional tool for the determination of system risk.

This paper has presented a number of methods that can be used to gauge the expected failure events in a system of computer hosts.

References

1. Benveniste, A. and Jacod, J. (1973). Systèmes de Lévy des processus de Markov. *Invent. Math.* **21** 183--198. Mathematical Reviews

[6] http://www.dshield.org/reports.html

2. Brémaud, P. (1981). *Point Processes and Queues: Martingale Dynamics*. Springer, New York. Mathematical Reviews
3. Chakrabarty, A. and Guo, X. (2007). A note on optimal stopping times with filtration expansion. Preprint, Univ. California, Berkeley.
4. Corcuera, J. M., Imkeller, P., Kohatsu-Higa, A. and Nualart, D. (2004). Additional utility of insiders with imperfect dynamic information. *Finance and Stochastics* **8** 437--450.
5. Dellacherie, C. and Meyer, P. A. (1982). *Probabilities and Potential. B*. North-Holland, Amsterdam.
6. Dempster, A.P. and Laird, N.M. and Rubin, D.B. (1977).Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, B39.
7. Elliott, R. J., Jeanblanc, M. and Yor, M. (2000). On models of default risk. *Math. Finance* **10** 179--195.
8. Grandell, J.: Aspects of Risk Theory. Springer, New York (1991)
9. Guo, X., Jarrow, R. and Zeng, Y. (2005). Modeling the recovery rate in a reduced form model. Preprint, Cornell Univ.
10. He, S. W., Wang, J. G. and Yan, J. A. (1992). *Semimartingale Theory and Stochastic Calculus*. Science Press, Beijing.
11. Hong Zhu, Yanlong Zhang, Qingning Huo, Sue Greenwood, "Application of Hazard Analysis to Software Quality Modelling," Computer Software and Applications Conference, Annual International, p. 139, 26th Annual International Computer Software and Applications Conference, 2002
12. Ikeda, N. and Watanabe, S. (1962). On some relations between the harmonic measure and the Lévy measure for a certain class of Markov processes. *J. Math. Kyoto Univ.* **2** 79--95.
13. Jacod, J. (1975). Multivariate point processes: Predictable projection, Radon--Nikodým derivatives, representation of martingales. *Z. Wahrsch. Verw. Gebiete* **31** 235--253.
14. Jeanblanc, M. and Valchev, S. (2005). Partial information and hazard process. *Int. J. Theor. Appl. Finance* **8** 807--838.
15. Kay, R. "Proportional Hazard Regression Models and the Analysis of Censored Survival Data" *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, Vol. 26, No. 3 (1977), pp. 227-237 Blackwell Publishing.
16. Marti, K. (2008) "Computation of probabilities of survival/failure of technical, economic systems/structures by means of piecewise linearization of the performance function",Structural and Multidisciplinary Optimization, Vol35/3, Pp 225 - 244.
17. Newman, M.E.J., Strogatz, S.H., and Watts, D.J. (2001) "Random graphs with arbitrary degree distributions and their applications". Phys. Rev. E 64. paper 026118
18. Revuz, D. and Yor, M. (1999). *Continuous Martingale and Brownian Motion*, 3rd ed. Springer, Berlin.